10/009571

3/prts

JC13 Rec'd PCT/PTO 12 DEC 2001

Novel calpains and the use thereof

The invention relates to a novel mammalian calpain CAPN11, to its synthesis and to its use.

calpains are a superfamily of related proteins, some of which have been demonstrated to act as calcium-dependent cysteine proteases. Eight different calpains have been identified in mammals.

calpains form a family of intracellular calcium-dependent cysteine proteases. An increasing number of mammalian calpain homologs has been identified, and the individual members can be divided into four classes on the basis of the physical structure and the predicted properties. Class A, the "classical calpains" CAPN1, CAPN2, CAPN3 (p94); CAPN8 (nCL-2) and CAPN9 (nCL-4) probably all have protease activity and are Ca^{2+}-dependent. They consist of a subunit of variable size (80 kDa) and an invariant small subunit (30 kDa). Class B and D calpains, CAPN5 (6, 15) and CAPN7 (8) have protease activity but are very probably Ca^{2+}-independent; the class C calpain, CAPN6, probably has no protease activity. The calpains can also be divided into categories on the basis of their expression patterns, with CAPN3, CAPN6, CAPN8 and CAPN9 showing some tissue specificity. The function of the calpains is unknown although they have been connected with a large number of physiological processes and pathological states (review in reference 17). To elucidate their function and evolutionary history it is necessary to identify the entire spectrum of calpain family members.

The invention relates to the novel polypeptide CAPN11 having the amino acid sequence disclosed in SEQ. ID No. 2.

The novel calpain protein CAPN11 has the properties typical of calpains including potential protease and calcium-binding domains. It displays a greatly restricted tissue distribution and is expressed mainly in the testis. With the aid of radiation hybrid mapping, the gene has been localized to chromosome 6 in a region which is assigned to p12. Phylogenetic analysis reveals that CAPN11 in mammals is most closely related to CAPN1 and CAPN2.

However, the predicted CAPN11 sequence has the greatest homology with the chicken calpain μ/m among the available calpain sequences. Thus, CAPN11 may be the human ortholog of μ/m calpain. The discovery of this novel calpain emphasizes the complexity of

the calpain family, the members of which can be distinguished on the basis of the protease activity, calcium dependence and tissue expression.

The cDNA nucleotide sequence of the CAPN11 gene contains 2338 nucleotides (SEQ ID No. 1). The cDNA sequence is derived from a single mRNA by successful amplification of the complete presumed coding region from human testis cDNA by means of flanking primers. Several cDNA clones have been completely sequenced in order to exclude any PCR artefacts.

There is a large open reading frame which encodes a protein having 702 amino acids (Mr 80 kDa) (Fig. 1). The amino acid sequence resembles the large subunit of members of the calpain family (Fig. 1). The protein can be divided into the four calpain-typical domains. Domain II displays the properties of a protease domain, and the predicted amino acid sequence has the three amino acid residues (Cys102, His259 and Asn283) which are part of the active site of cysteine proteases (2). The amino acid sequence of all five Ca^{2+}-binding sequences described for CAPN2 (4, 12) are conserved to a certain extent (Fig. 1). This protein thus probably has protease and calcium-binding properties. Comparison of the predicted amino acid sequence with those of all other calpains revealed the greatest sequence homology (57.5%) with the chicken calpain μ/m. In mammalian calpains, the greatest similarity was with human CAPN1 (54.3% homology). The least similar human calpain, with only 18.7% homology, was CAPN6. The gene coresponding to this cDNA has been designated CAPN11 by the Human Gene Nomenclature Committee.

The complete amino acid sequence of all identified human calpains has undergone phylogenetic analysis. The results make it possible to classify human calpains into four main evolutionary groups (Fig. 2). The first group is represented by CAPN5, CAPN6, CAPN7 and CAPN8, the second by CAPN1 and CAPN2, the third group by CAPN3 and CAPN9, and the fourth comprises CAPN11. The phylogenetic analysis thus suggests that CAPN11 represents a separate calpain subfamily.

The expression of CAPN11 in human tissues was investigated by Northern and RNA dot-blot analyses. Of the 50 tissue RNAs investigated, the CAPN11 mRNA was expressed most strongly in the testis (Fig. 3A). The specificity of this signal was confirmed by Northern blot analysis and corresponded to an mRNA about 3 kb in size (Fig. 3D). Much weaker signals were detected in the thymus and in the mammary gland. The significance of this finding is, however, unclear because further investigation of thymus RNA by

Northern blot analysis produced no signal, despite long exposure times (Fig. 3D and data not shown). A possible explanation would be that this weak signal is attributable to cross-hybridization with related mRNAs. The testis is thus the main expression site of CAPN11, although we are unable to preclude the possibility that the gene is expressed in other tissues which were not investigated.

We have determined on which chromosome the human CAPN11 gene is located. We have assigned by means of PCR with primers which are specific for the human CAPN11 nucleotide sequence, and with a somatic human/rodent cell hybrid mapping panel, the gene to chromosome 6 (Correll Cell Repositories). With the aid of radiation hybrid mapping using the medium-resolution Stanford G3 panel (Research Genetics Inc.) and the database at the Stanford Human Genome Center (shgc-www.stanford.edu), the gene was located 5 centiray away from the marker SHGC-32834 on this chromosome (LOD score 12.87). This marker is located in the interval between the microsatellite markers D6S1616 (59.6 cM) and D6S427 (73.9 cM) (7), and a marker in this interval, D6S269, has been cytogenetically assigned to 6p12 (5). Thus, CAPN11 is located on chromosome 6 in the vicinity of p12. No other calpain gene has been located on this chromosome.

The chicken μ/m calpain was the first member of the calpain family to be cloned (16). It was originally referred to as m calpain, but was reclassified after identification of other chicken calpains which are very probably orthologous to the mammalian μ and m calpains (18). However, no mammalian μ/m calpain has yet been determined. Since, however, CAPN11 displays greater homology with chicken μ/m calpain than with other mammalian calpains, it may be the ortholog thereof.

There are to date 5 calpains displaying a certain degree of tissue specificity - CAPN3 (skeletal muscle), CAPN6 (placenta), CAPN8 (possibly smooth muscle), CAPN9 (stomach and small intestine) and CAPN11 (testis). Numerous proteases have been identified in the testis, and it is assumed that they are involved in processes such as tissue reorganization (20), regulation of spermatogenesis (14), penetration of the Zona pellucida by sperm (10) and fertility (13). However, many of these activities are dependent on secreted proteases, and CAPN11 probably has an intracellular localization. In the testis, it might be involved in processes in which calpains in other tissues are involved, such as germ cell apoptosis (3) or regulation of testis-specific transcription factors.

A further aspect of this invention relates to the use of the polypeptide CAPN11 for identifying substances which are able to inhibit the enzymatic activity of this polypeptide, called calpain inhibitors, in particular those calpain inhibitors which are selective for CAPN11. Selectivity means that such calpain inhibitors inhibit the activity of CAPN11 more than the activity of the other calpains mentioned above, in particular preferably at least 10 times, and more preferably 25 times, more. The enzymic activity of CAPN11 is a Ca-dependent protease activity.

A further aspect of the invention relates to a method for identifying compounds which inhibit the enzymic activity of a polypeptide as claimed in claim 1, comprising:

(a) comparing the extent of the enzymatic activity of CAPN11 in the presence of the compound with the extent of the enzymatic activity of CAPN11 in the absence of the compound and

(b) selecting compounds which alter the extent of the enzymatic activity of CAPN11 compared with the enzymatic activity of CAPN11 in the absence of the compound.

The inhibiting compounds identified by the method mentioned above are suitable for the treatment of disorders associated with or linked to a nonphysiologically elevated CAPN11 activity, such as infertility in men.

The dosage and the treatment regimen for these inhibitors must be determined by routine methods known for other protease inhibitors.

REFERENCES

1. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment sequencing tool. J Mol. Biol. 215:403-410.
2. Berti, P. J and Storer, A. C. (1995). Alignment/Phylogeny of the papain superfamily of cysteine proteases. J. Mol. Biol. 246: 273-283.
3. Billig, H., Chun S.Y., Eisenhauer K. and Hsueh, A.J. (1996). Gonadal cell apoptosis. Hormone-regulated cell demise. Hum. Reprod. Update 2:103-117.
4. Blanchard, H., Grochulski, P., Li, Y., Arthur, J.S.C., Davies, P.L., Elce, J.S. and Cygler, M. (1997). Structure of a Ca^{2+}-binding domain reveals a novel EF-hand and Ca^{2+}-induced conformational changes. Nature Structural Biology 4: 532-538.

5. Bray-Ward, P., Menninger, J., Lieman, T., Desai, T., Mokady, N., Banks, A. and Ward, D.C. (1996). Integration of the cytogenetic, genetic and physical maps of the human genome by FISH mapping of CEPH YAC clones. Genomics 32: 1-14.

6. Dear, N., Matena, K., Vingron, M. and Boehm, T. (1997). A new subfamily of vertebrate calpains lacking a calmodulin-like domain: Implications for calpain regulation and evolution. Genomics 45:175-184.

7. Deloukas, P. et al. (1998). A physical map of 30,000 human genes. Science 282: 744-746.

8. Franz, T., Vingron, N., Boehm, T. and Dear, T.N. (1999). Capn7: A highly divergent vertebrate calpain with a novel C-terminal domain. Mamm. Genome, 10:318-321.

9. Frohman, M. A., Dush, M. K. and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA 85: 8998-9002.

10. Kohno, N., Yamagata, K., Yamada, S., Kasiwabara, S., Sakai, Y. and Baba, T. (1998). Two novel testicular serine proteases, TESP1 and TESP2, are present in the mouse sperm. acrosome. Biochem. Biophys. Res. Commun. 245:658-665.

11. Kozak, M. (1996). Interpreting cDNA sequences: some insights from studies on translation. Mamm. Genome 7: 563-574.

12. Lin, G.-D., Chattopadhyay, D., Maki, M., Wang, K.K.W., Carson, M., Jin, L., Yuen, P.-W., Takano, E., Hatanaka, M., DeLucas, L.J. and Narayana, S.V.L. (1997). Crystal structure of calcium bound domain VI of calpain at 1.9 A_i resolution and its role in enzyme assembly, regulation, and inhibitor binding. Nature Structural Biology 4: 539-547.

13. Mbikay, M., Tadros, H., Ishida, N., Lerner, C.P., De Lamirande, E., Chen, A., El-Alfy, M., Clermont, Y., Seidah, N.G., Chrétien, N., Gagnon, C. and Simpson, E.M. (1997). Impaired fertility in mice deficient for the testicular germ-cell protease PC4. Proc. Natl. Acad. Sci. USA 94:6842-6846.

14. Monsees, T.K., Gornig, N., Schill, W.B. and Miska, W. (1998). Possible involvement of proteases in the regulation of spermatogenesis. Andrologia 30:185-191.

15. Mugita, N., Kimura, Y., Ogawa, M., Saya, H. and Nakao, M. (1997), Identification of a novel, tissue-specific calpain htra-3; a human homologue of the Caenorhabditis elegans sex determination gene. Biochem. Biophys. Res. Comm. 239:845-850.

16. Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M. and Suzuki, K. (1984). Evolutionary origin of a calcium-dependent protease by fusion of genes for a thiol protease and a calcium-binding protein? Nature 312: 566-570.

17. Ono, Y., Sorimachi, H. and Suzuki, K. (1998). Structure and physiology of calpain, an enigmatic protease. Biochem. Biophys. Res. Commun. 245:289-294.
18. Sorimachi, H., Tsukahara, T., Okada-Ban, M., Sugita, H., Ishiura, S. and Suzuki, K. (1995). Identification of a third ubiquitous calpain species - chicken muscle expresses four distinct calpains. Biochim. Biophys. Acta 1261: 381-393.
19. Thompson J.D., Higgins D.G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22:4673-4680.
20. Tohonen, V., Osterlund, C. and Nordqvist, K. (1998). Testatin: A cystatin-related gene expressed during early testis development. Proc Natl. Acad. Sci. USA 95:14208-14213.

FIGURES

FIG. 1. Grouping of the predicted amino acid sequence of CAPN11 with other human calpains. The multiple grouping of the amino acid sequences was carried out using CLUSTAL W (19). The presumptive starting methionine for CAPN11 (GGAatgG) corresponds to the minimal consensus sequence for the translation starting site (RNNatgG, where R is a purine, Ref. 11). Amino acids in the other proteins which are identical to those of CAPN11 are shaded. Dashes indicate gaps which have been introduced to maximize the alignment. Arrowheads point to the three conserved amino acids which are part of the active site of the calpains. The potential EF-hand calcium-binding domains of CAPN8 are underlined and numbered consecutively. The arbitrary domains of calpain are indicated. The sequences for CAPN4 and CAPN7, which have been identified only in rat and mouse respectively, are not shown. The published CAPN6 sequence has not been included because it is only a partial sequence. The alternative names and accession numbers for the grouped calpains are indicated in the legend for Fig. 2.

FIG. 2. Rootless phylogenetic tree of the family of the large subunit of human calpains. The analysis was carried out with the PAUP program, and the tree was arranged using CLUSTREE from the HUSAR server of the Deutsche Krebsforschungszentrums, Heidelberg (www.dkfz-heidelberg.de). The lengths of the horizontal lines are proportional to the derived phylogenetic distances; the vertical lines have no significance. 1000 bootstrapping repetitions were carried out, and the values are shown at the internal nodes. The CAPN7 sequence is derived from the mouse because only little of the human nucleotide and protein sequence is available. The human ortholog does in fact exist (see Ref. 8) so that the use of the

mouse sequence for this comparison is justified. The partial human CAPN8 sequence is the predicted translation of the EST clone AA026030 (Hillier et al., 1995, The WashU-Merck EST project, unpublished results). An amino acid translation of this clone shows great similarity with the rat CAPN8 sequence. No bootstrapping was carried out with this sequence because it is much shorter than the others. Thus no sensible comparison of the bootstrapping value with the values from comparisons of full-length sequences is possible. The nomenclature specified by the Human Gene Nomenclature Committee is used. The previous names for the various calpains are: CAPN1 - m calpain, CAPN2 - m calpain, CAPN3 - p94, nCl-1, CAPN8, nCL-2, CAPN9, nCL-4. The EMBL accession numbers for the calpain sequences used are: CAPN1 (P17655), CAPN2 (P07384), CAPN3 (P20807), CAPN5 (Y10656), CAPN6 (Y12582), CAPN7 (AJ012475) and CAPN9 (AF022799).

FIG. 3. Expression of CAPN11. A ^{32}P-labeled DNA probe with an 800 base pair segment of the coding sequence of the human CAPN11 cDNA was hybridized on a master blot (A), a nylon filter with dot-blots of RNAs from 50 different human tissues or a Clontech multiple tissue Northern blot (D). The filters were washed with high stringency (6x SSC, 65°C). The exact site of the various RNAs on the dot-blot filter is shown diagrammatically (C). The RNAs on the Northern blot are indicated over the corresponding lanes. Dot-blot and Northern blot were rehybridized with DNA probes for human ubiquitin (B) and β-actin in order to determine the amounts of poly(A+) RNA loaded. The sites of the size markers (in kilobases) are indicated for the Northern blot. The exposure times were: A, 72 h; B, 24 h; D, 48 h, PBL = peripheral blood leukocytes.